LEGIBILITY NOTICE

A major purpose of the Technical Information Center is to provide the broadest dissemination possible of information contained in DOE's Research and Development Reports to business, industry, the academic community, and federal, state and local governments.

Although a small portion of this report is not reproducible, it is being made available to expedite the availability of information on the research discussed herein.

LA-UR -87-3115

CONF-8710151--1

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

LA-UR--87-3115

DE88 000537

TITLE: CONTROLLING ROBOTS WITH SPOKEN COMMANDS

AUTHOR(S) Tony Beugelsdijk, MEE-9
Patrick Phelan, MST-12

SUBMITTED TO Fifth International Symposium on Laboratory Robotics
Zymark Corp.
Zymark Center
Hopkinton, MA 01748

DISCLAIMER

MASTER

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

CONTROLLING ROBOTS WITH SPOKEN COMMANDS

Tony Beugelsdijk and Patrick Phelan
Los Alamos National Laboratory
P.O. Box 1663
Los Alamos, NM 87544

ABSTRACT

A robotic system for handling radioactive materials has been developed at Los Alamos National Laboratory. Because of safety considerations, the robot must be under the control of a human operator continuously. In this paper we describe the implementation of a voice-recognition system that makes such control possible, yet permits the robot to perform pre-programmed manipulations without the operator's intervention. We also describe the training given both the operator and the voice recognition-system, as well as practical problems encountered during routine operation.

A speech synthesis unit connected to the robot's control computer provides audible feedback to the operator. Thus, when a task is completed or if an emergency develops, the computer provides an appropriate spoken message. Implementation and operation of this commercially available hardware are discussed.

INTRODUCTION

At our Laboratory, we find it desirable to use robots for numerous processes involving remote manipulation of materials. However, many of the processes are not repetitive, or they treat hazardous material, so a human operator must be in control at all times. Accordingly, it is extremely useful to let the operator control the robot by

speaking commands into a microphone. The commands initiate pre-programmed procedures as soon as they are spoken.

Speaking the commands is more convenient than entering them at the keyboard because the operator's hands remain free for other tasks, and he or she may walk about. Speaking commands is also more convenient than using a "teach pendant" because the pre-programmed movements are executed faster and more accurately. Finally, because the operator needs little knowledge of how the robot works or of its programming language, he or she can operate it successfully with little training.

EXPERIMENTAL

Components

The major parts of our system are depicted in Fig. 1. The laboratory robot (Model 9000 MasterLab™ system, Perkin-Elmer Corp., Ridgefield, CT) is under direct control of a microcomputer (Personal Computer XT™, International Business Machines, Boca Raton, FL) linked to a control unit. A voice-recognition system interprets commands spoken into the microphone and passes them on to the computer. The voice-recognition system is either a special keyboard with built-in electronics (Model KB 5152V, KeyTronic Corp., Spokane, WA) or an interface card installed in the computer chassis (Vocalink™ model SRB-LC, Interstate Voice Products, Orange, CA).

The speech-synthesis unit (Echo-PC™, Street Electronics Corp., Carpinteria, CA) is connected to an RS-232 serial port as if it were a printer.

Installing the Voice-Recognition System

The first step is to develop the desired robotic procedures, storing each sequence of robot moves as a computer program that can be initiated by entering the program's name at the keyboard. Each procedure must be tested thoroughly before being put into use.

The second step is to install the voice-recognition system in the computer and connect the microphone. For a keyboard system with built-in electronics, installation means simply plugging its cable into the personal computer. If the system is an interface card, however, some care must be taken to ensure that its I/O bus address does not conflict with the addresses of other interface cards that are already present. If there is a conflict, the address can be changed by switches mounted on the card.

The third step is to enter a list of commands into the computer. Using software supplied by the manufacturer of the voice-recognition system, the user assigns each command a character string that is the name of one pre-programmed procedure.

Finally, each operator must train the voice-recognition system to recognize his or her voice. Software provided by the manufacturer is again used. To do the training, the operator speaks each command into the microphone when prompted by the computer. The words are repeated three to five times until the voice-recognition system has a pattern it considers typical of the individual's voice.

Installing the Speech-Synthesis Unit

All the necessary electronics and programming already reside in this type of speech-synthesis unit. Dip switches on the bottom of the unit permit the user to select the baud rate over the range 75 to 9600.

How the Voice-Recognition System Works

First, each word spoken into the microphone is compared with those in the operator's previously stored vocabulary list. The voice-recognition system compares the spoken word's digitized pattern with the patterns of the vocabulary words.

When a satisfactory match is found between the spoken command and a vocabulary word, the computer recalls the character string associated with the vocabulary word and sends the string to the robot controller as if the string had been entered at the keyboard. The robot acts on the string by performing the pre-programmed procedure in the usual manner.

How the Speech-Synthesis Unit Works

PRINT commands included in each pre-programmed robotic procedure send selected character strings to the unit as if it were a printer. When the speech-synthesis unit receives a character string, the unit processes it according to nearly 400 rules of English pronunciation, then provides an audible word or words through its built-in speaker.

RESULTS

Voice-Recognition System

After their initial training, operators found the robot easy to operate by speaking commands. The voice-recognition system correctly recognized 90 percent of the words the first time they were spoken. Usually, whenever the command was not recognized the first time, the system and the robot did not respond at all; but they did respond correctly when the command was spoken a second time. With proper adjustment of the pattern-recognition criteria (as discussed below), the frequency with which the system misinterpreted commands was reduced to less than 1 percent.

The systems we used are so sensitive to subtle differences in pronunciation that it was necessary for each operator to record a vocabulary of commands and train the system to recognize his or her unique way of pronouncing them. In general, when an operator tried to use the vocabulary that

had been recorded by another, the recognition frequency was less than 25 percent.

Speech-Synthesis Unit

Users must exercise care in selecting the words to be output through the speech-synthesis unit. We did much trial-and-error testing to ensure that the words would be readily intelligible. Even so, about 25 percent of the words were unintelligible to inexperienced listeners the first time they heard them. Listeners' recognition rate grew to over 90 percent once they became accustomed to the unit's accent.

DISCUSSION

Our experience is limited to commercially available voice I/O hardware designed for use in an IBM personal computer. Therefore, we used spoken commands only with robotic systems that can be controlled by such a computer. Moreover, we used only systems that are among the least expensive. There are more-sophisticated systems available that have fewer limitations than the ones we employed.

Training Needed by Operators

It is helpful, of course, if the operators have some basic knowledge about how a personal computer works; for example, how to enter commands at the keyboard and how to check for loose cables. It is more important, however, that the operators be familiar with the robotic system. They need to

know what manipulations it has been programmed to carry out, and the command words used to initiate those manipulations.

Operators also need to be trained to use the voice-recognition system. For example, they need to know how to position the microphone, and they must practice saying the commands distinctly and with a noticeable pause between each command.

Choosing the Vocabulary of Commands

Perhaps the most difficult task involved in installing a voice-recognition system is preparing the vocabulary of command words. The commands must be easy for the operator to remember, of course, so they should be few in number. It also helps if the words describe the task the robot will perform. For example, the command "arm-up" is more suggestive of a procedure than "move".

Usually it is important to avoid words that sound similar, but the software provided with the system can often alleviate this problem. For example, after a vocabulary of commands has been recorded during the training period, the computer can test the digitized patterns and measure how similar they are. If two words are quite similar, the chance for confusion can be reduced by requiring the system to make a more stringent comparison between the pattern of the incoming word and the patterns of the problem words. However, there is a disadvantage to doing this; the system

may fail to recognize one of the vocabulary words if there is a slight variation in its pronunciation.

It is also a good idea to avoid one-syllable words, because the word's digitized pattern may be so simple that the system has trouble recognizing it. But, again, the software can alleviate this problem. The computer can test the digitized patterns in the vocabulary and predict how easily each command will be recognized. The computer can be instructed to be more lenient when comparing the incoming word with the simple words, if the user wishes. However, this will increase the chance that non-vocabulary words (and even extraneous noises) will be interpreted as commands..

The software provided by the manufacturer allows the vocabulary words to be grouped so that only a few words will be active at one time. This approach makes it easier for the system to recognize the commands successfully. An example of such a grouping is shown in Fig. 2. The words in the first group are activated when the operator speaks the command "move". Thus, the system will recognize only the words, "up", "down", "right" and "left" until the word "revolve" is spoken. Then only the words "around", "right" and "half" will be recognized. As illustrated by this example, the same word can appear in two different groups and can have different meanings depending upon which group is active.

SUMMARY

We successfully used commercially available hardware and software to control a robot by giving it spoken commands. Such control was augmented by audible responses from the robot through a speech-synthesis unit.

ACKNOWLEDGMENTS

The authors appreciate the assistance provided by their coworker, Chris Keddy, whose efforts contributed significantly to this study.

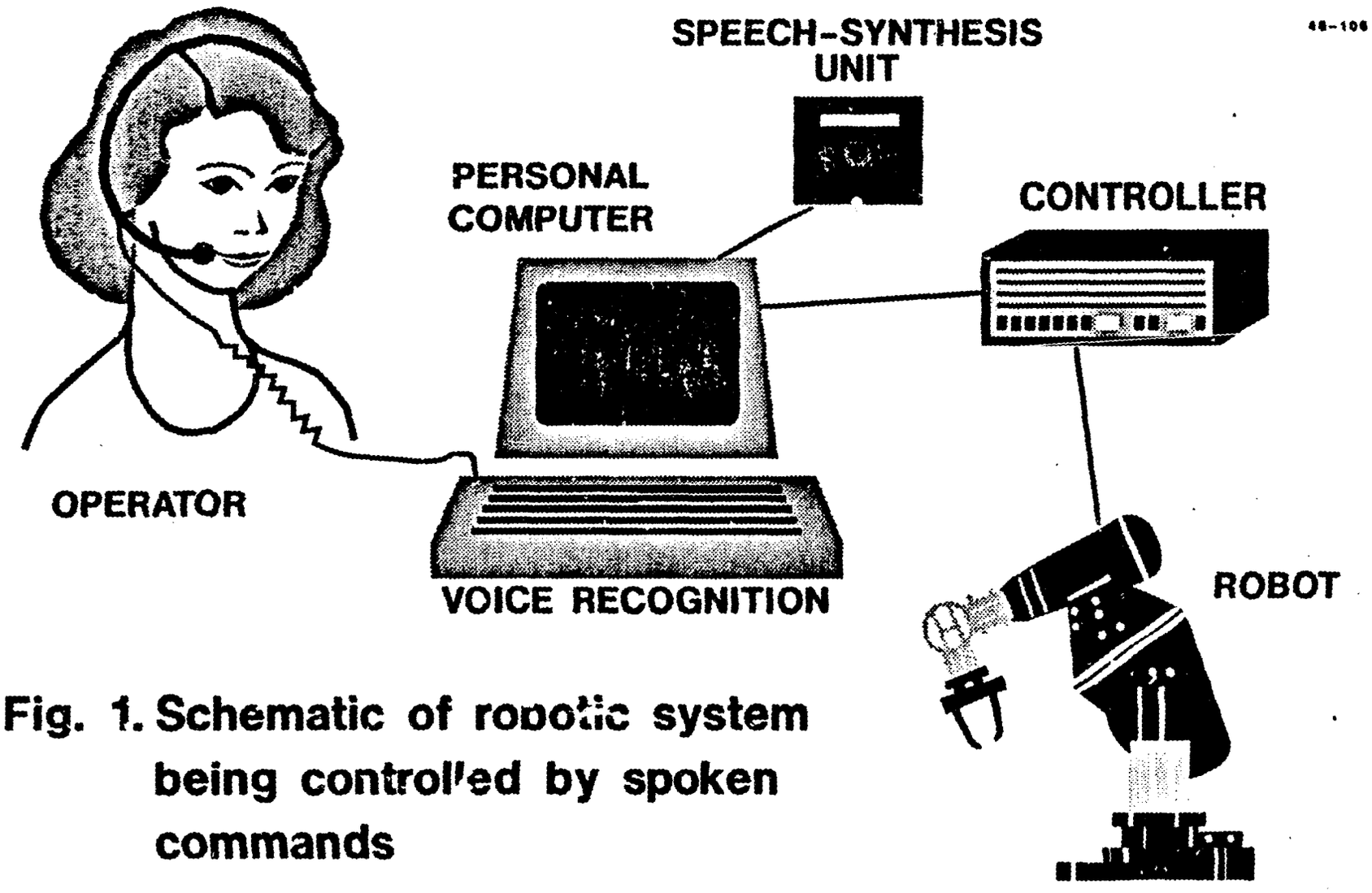

Fig. 1. Schematic of robotic system being controlled by spoken commands

Move

- **Up**
- **Down**
- **Right**
- **Left**

Revolve

- **Around**
- **Right**
- **Half**

Fig. 2. Command words grouped to improve recognition accuracy